Fig 1

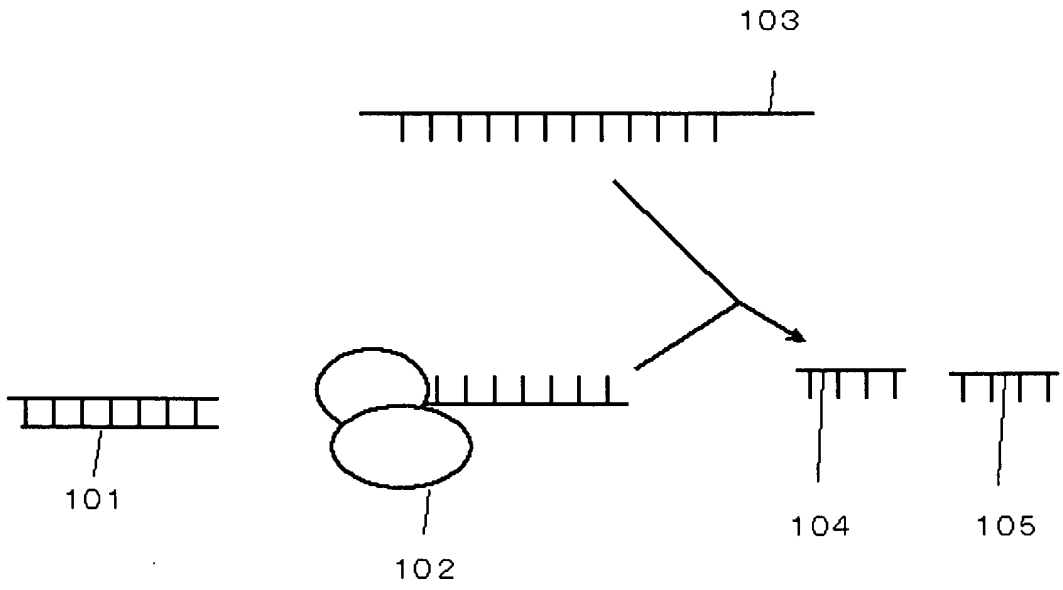


Fig. 2

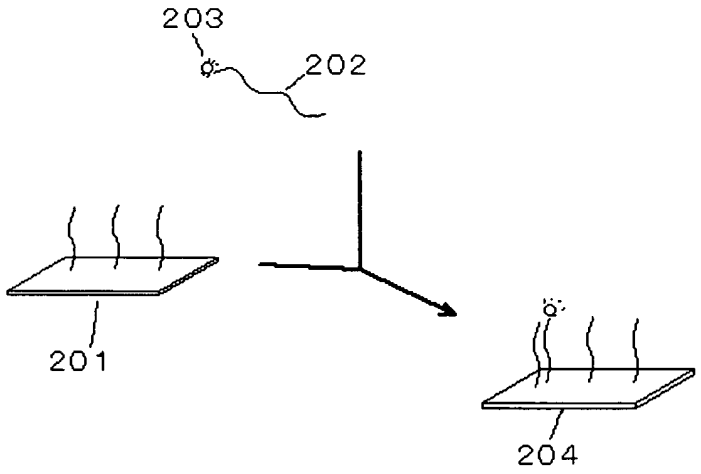


Fig. 3

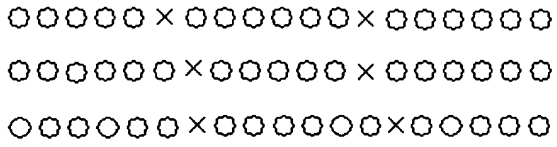


Fig. 4

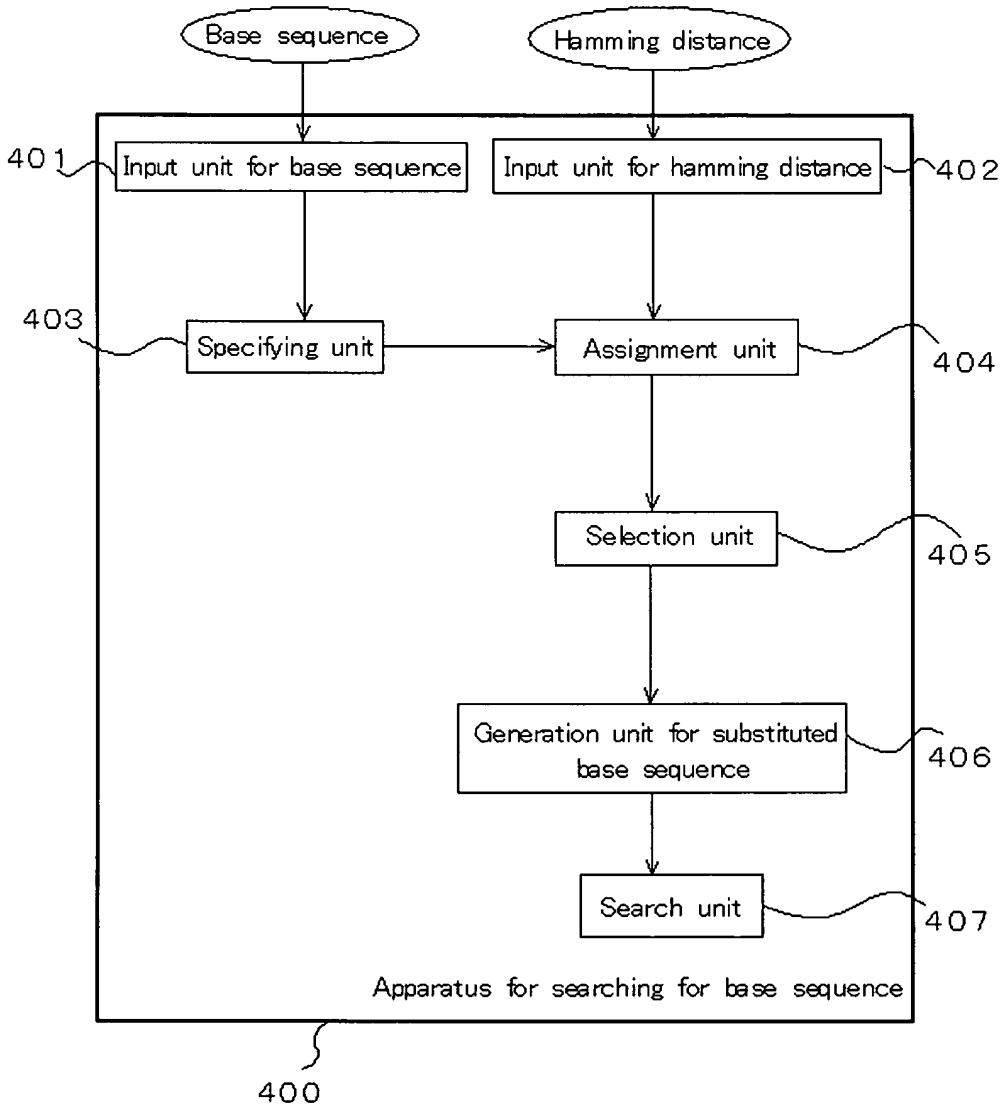
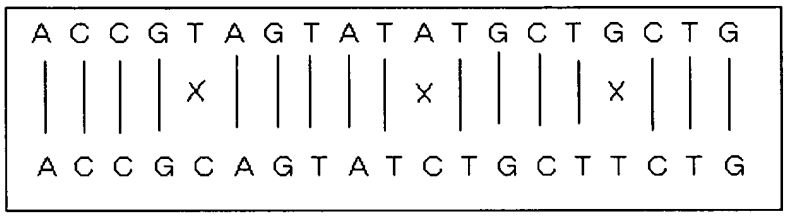

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
Search unit
407
Apparatus for searching for base sequence
400

Fig. 5



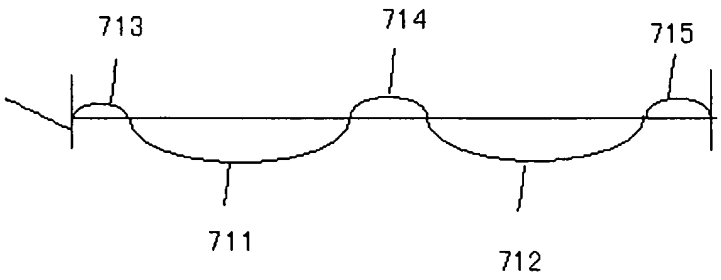
1454253A01092606

Fig. 6

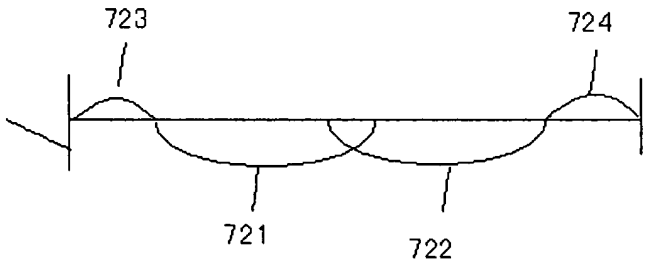
$$d_H(S,T) = \left| \{ i \mid S_i \neq T_i,\ i=1,2,\cdots,n \} \right|$$

Fig 7

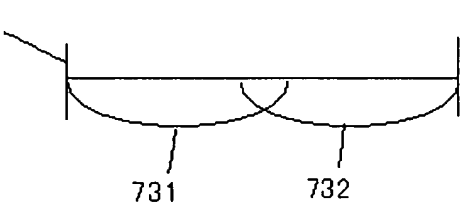
(1)


713
714
715
710
711
712

(2)


723
724
720
721
722

(3)


730
731
732

(4)

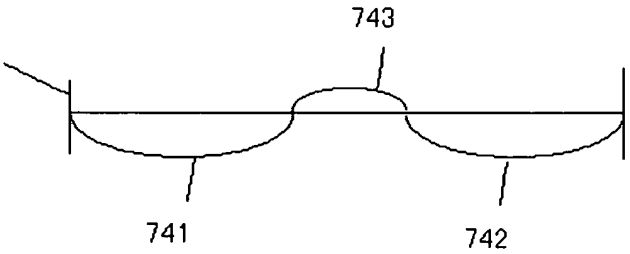
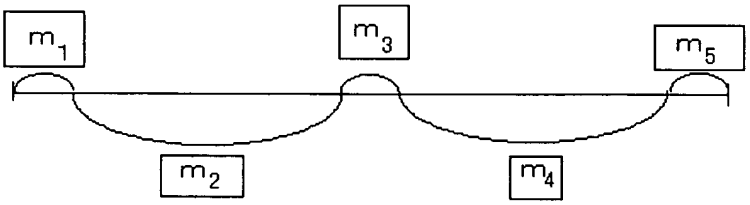

743
740
741
742

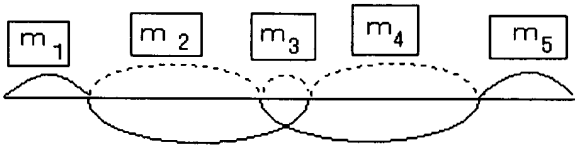
Fig. 8

(1)



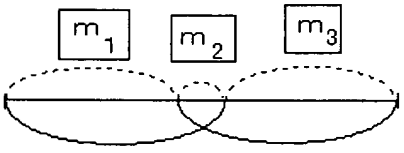
$m_1 + m_2 + m_3 + m_4 + m_5 =$ Inputted hamming distance

(2)



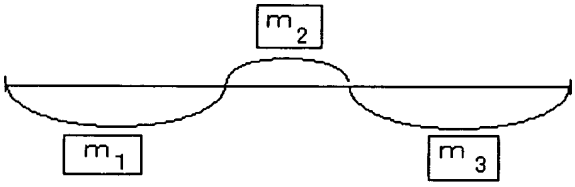
$m_1 + m_2 + m_3 + m_4 + m_5 =$ Inputted hamming distance

(3)



$m_1 + m_2 + m_3 =$ Inputted hamming distance

(4)



$m_1 + m_2 + m_3 =$ Inputted hamming distance

Fig 9

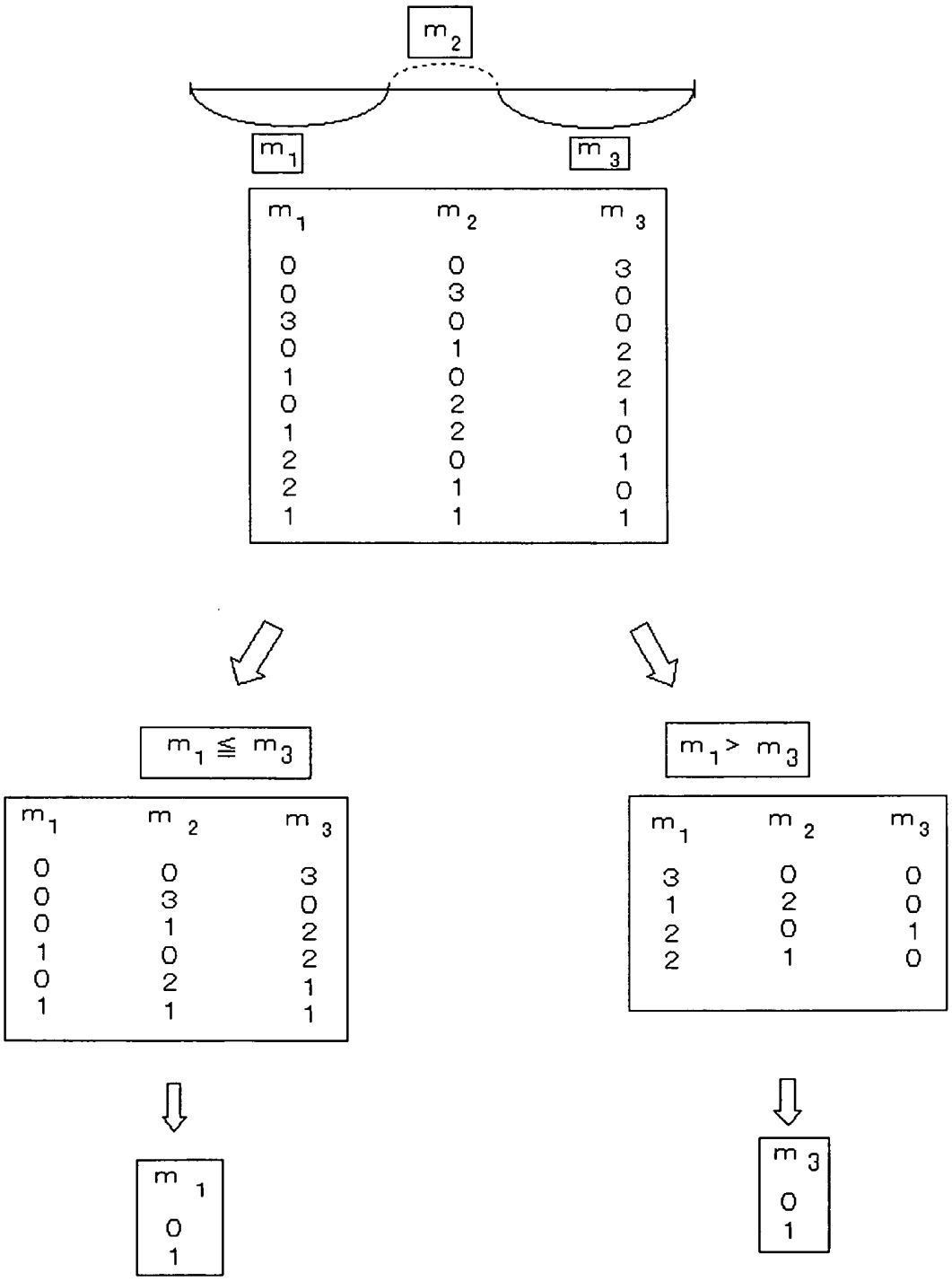
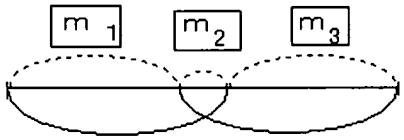

m2
m1
m3
m1 m2 m3
0 0 3
0 3 0
3 0 0
0 1 2
1 0 2
0 2 1
1 2 0
2 0 1
2 1 0
1 1 1
m1 ≦ m3
m1 m2 m3
0 0 3
0 3 0
0 1 2
1 0 2
0 2 1
1 1 1
m1
0
1
m1 > m3
m1 m2 m3
3 0 0
1 2 0
2 0 1
2 1 0
m3
0
1

Fig. 10



m_1	m_2	m_3
0	0	3
0	3	0
3	0	0
0	1	2
1	0	2
0	2	1
1	2	0
2	0	1
2	1	0
1	1	1

$m1 + m2 \leqq m2 + m3$

m_1	m_2	m_3
0	0	3
0	3	0
0	1	2
1	0	2
0	2	1
1	1	1

m1 + m2
0
1
2
3

$m1 + m2 > m2 + m3$

m_1	m_2	m_3
3	0	0
1	2	0
2	0	1
2	1	0

m2 + m3
0
1

Fig. 11

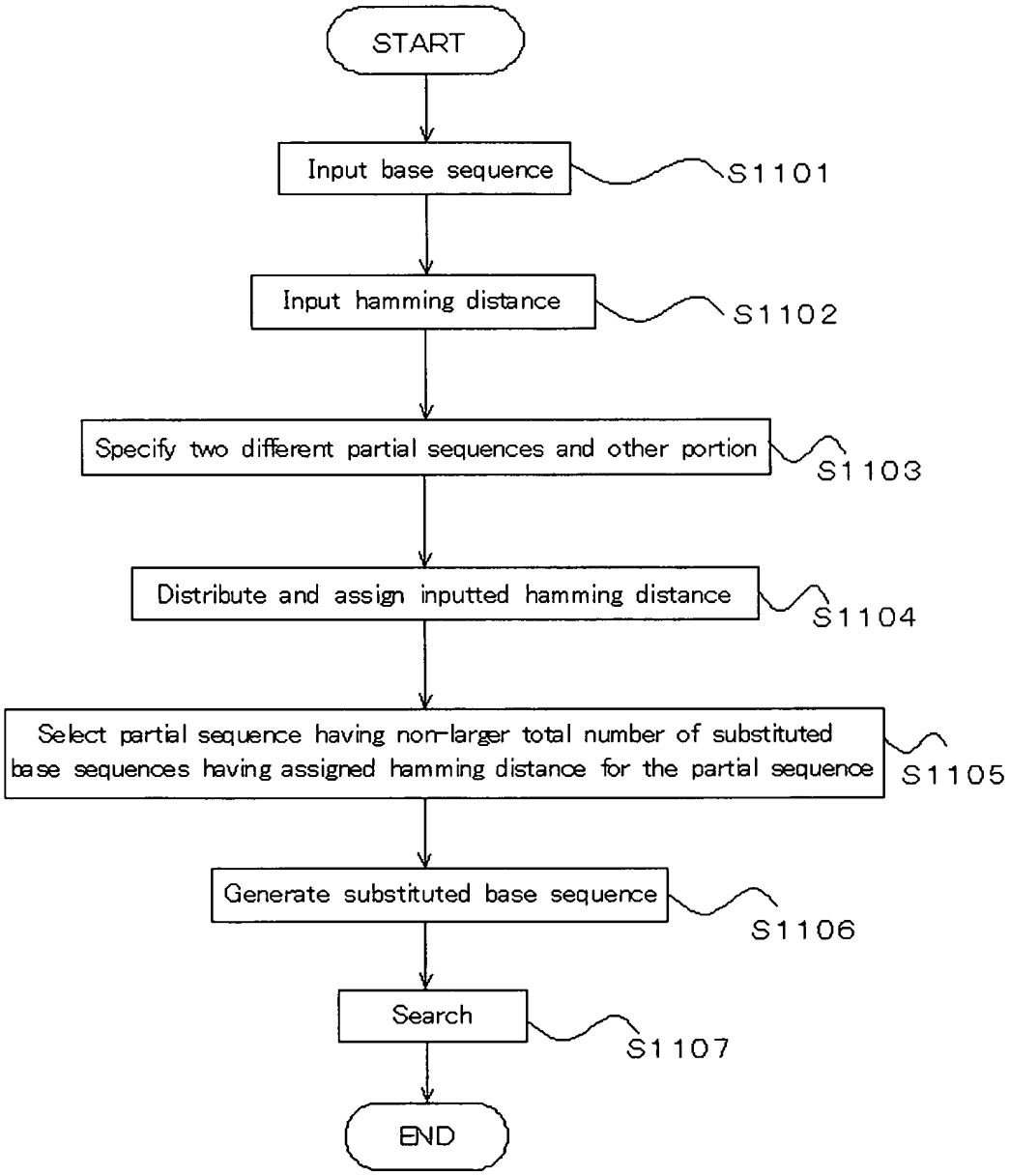

START
Input base sequence
S1101
Input hamming distance
S1102
Specify two different partial sequences and other portion
S1103
Distribute and assign inputted hamming distance
S1104
Select partial sequence having non-larger total number of substituted base sequences having assigned hamming distance for the partial sequence
S1105
Generate substituted base sequence
S1106
Search
S1107
END

Fig. 12

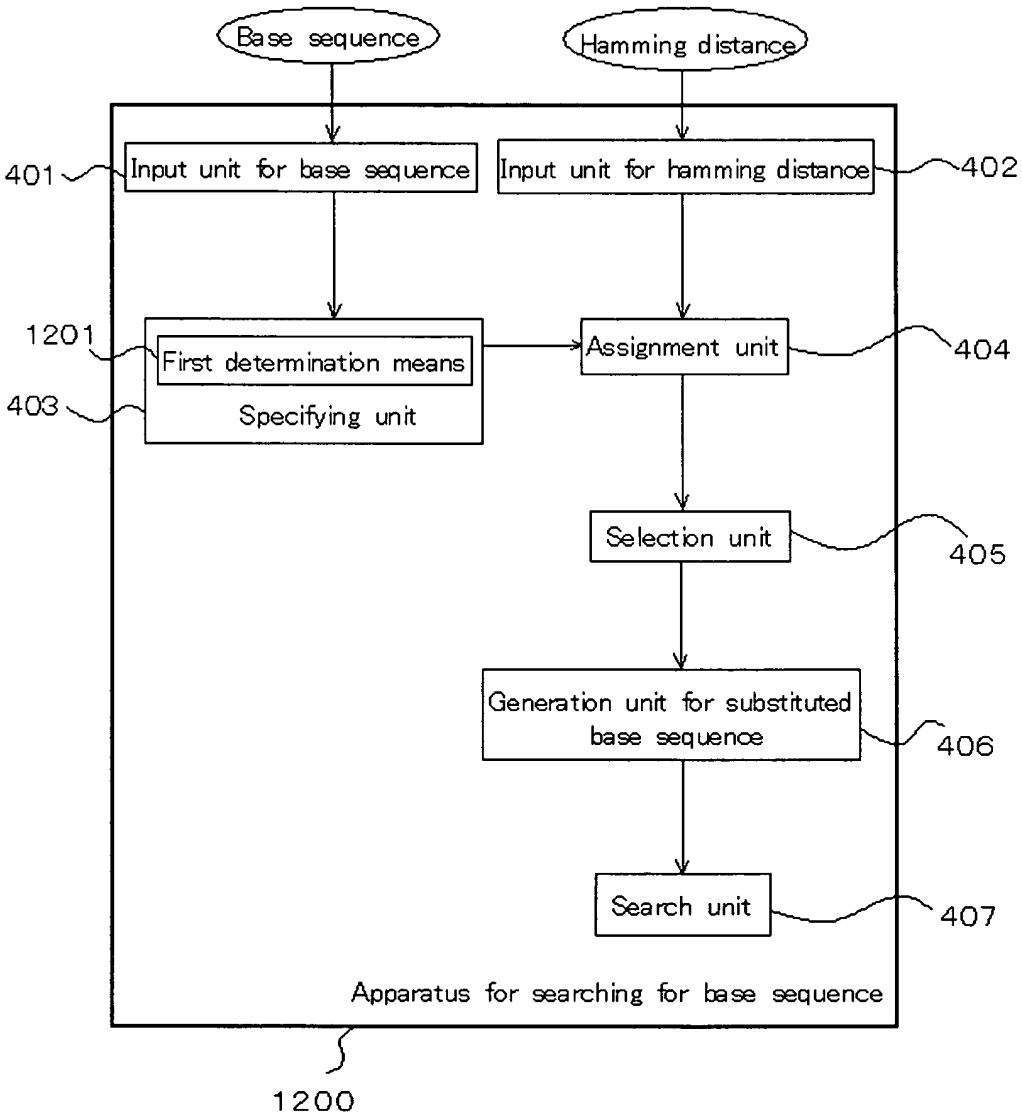

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
1201
First determination means
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
Search unit
407
Apparatus for searching for base sequence
1200

Fig. 13

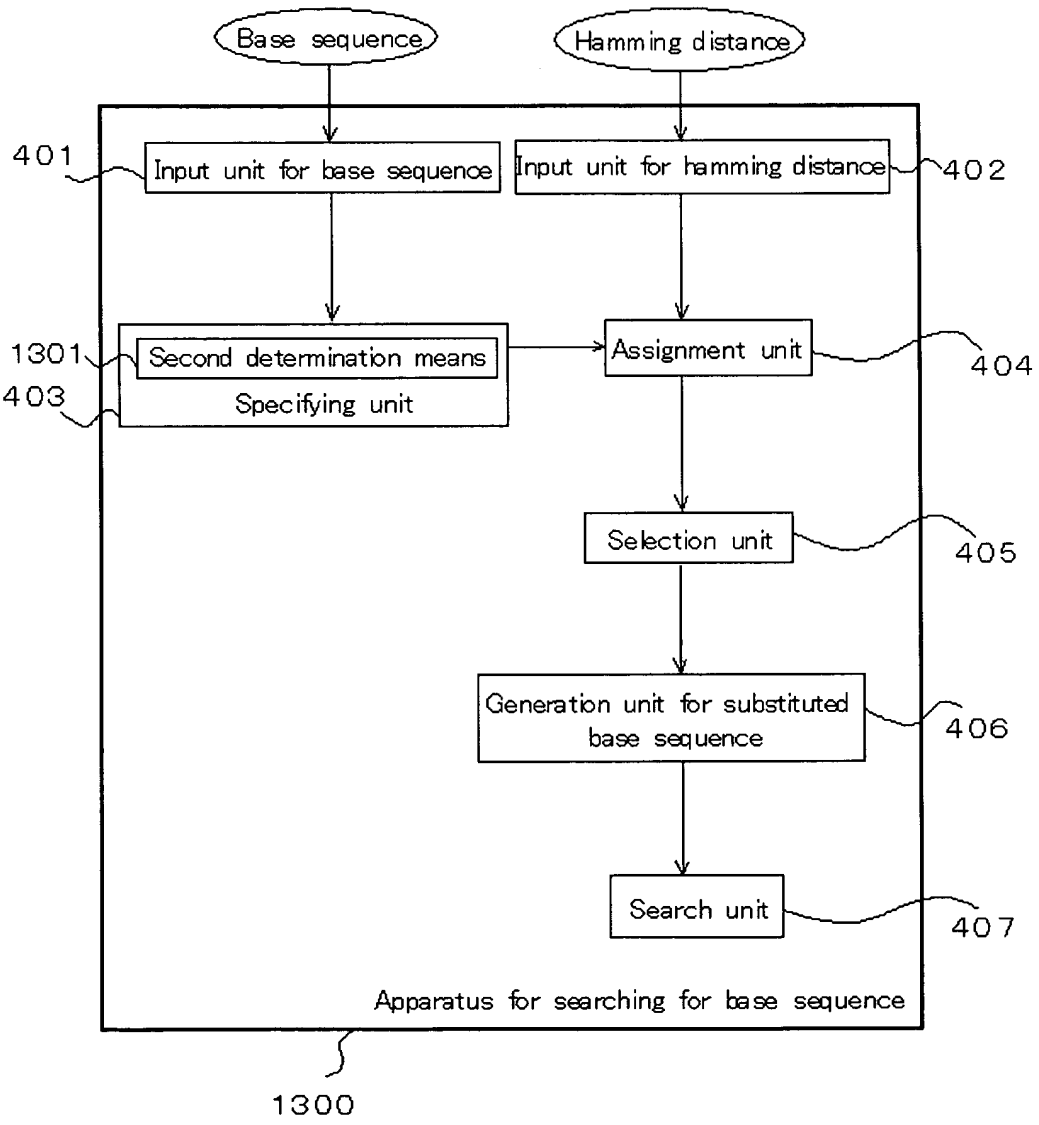

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
1301
Second determination means
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
Search unit
407
Apparatus for searching for base sequence
1300

Fig. 14

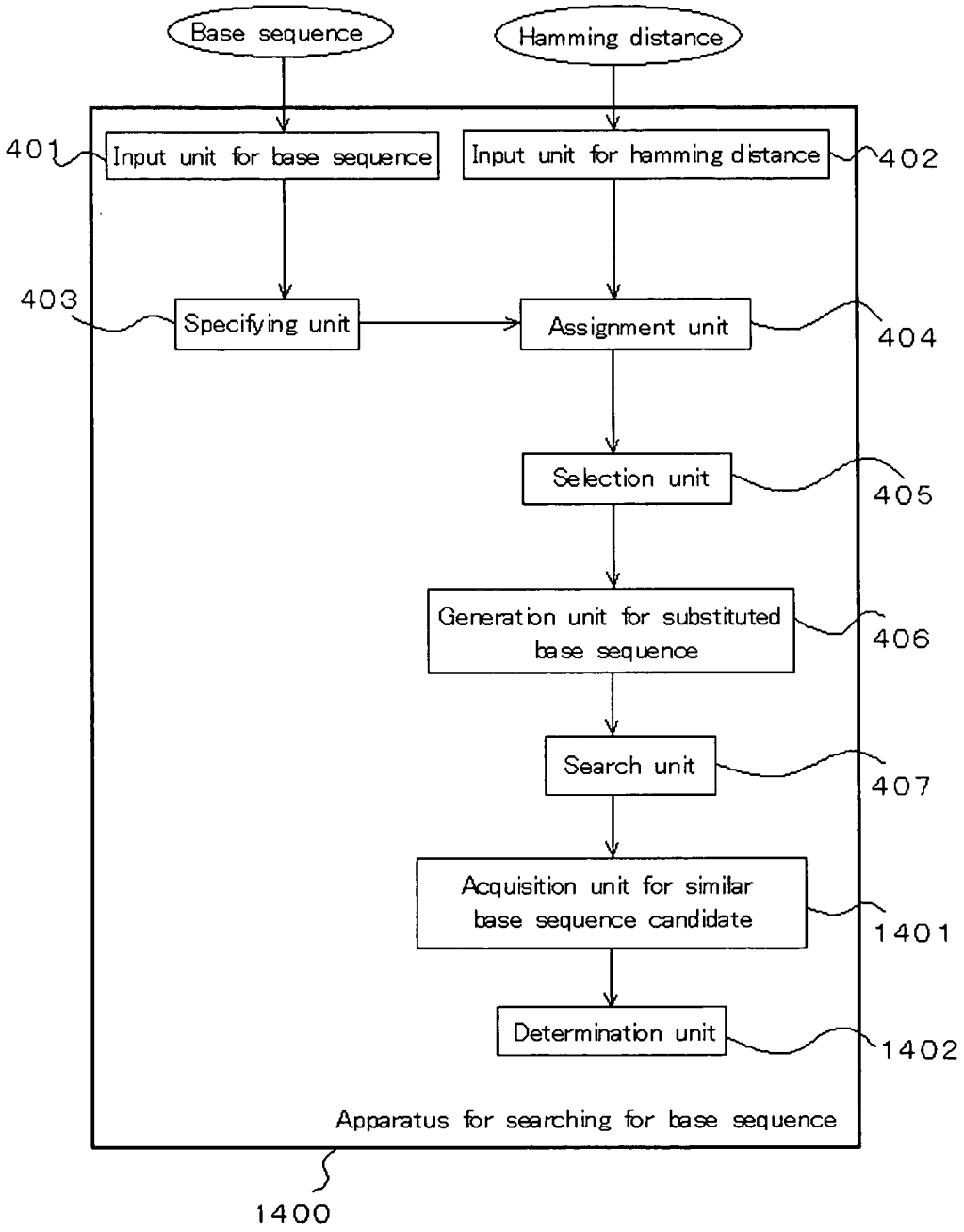

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
Search unit
407
Acquisition unit for similar base sequence candidate
1401
Determination unit
1402
Apparatus for searching for base sequence
1400

Fig 15

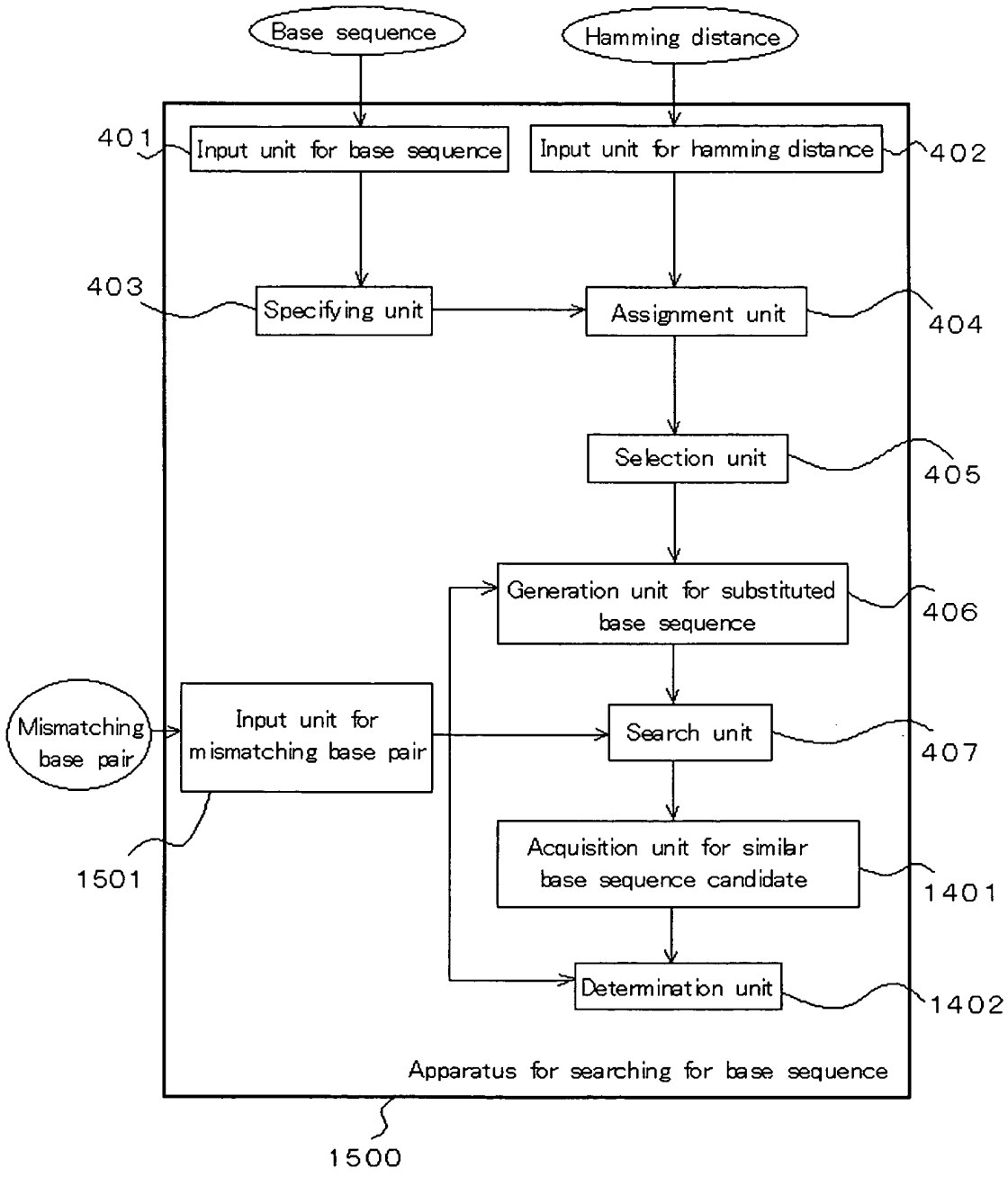

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
Mismatching base pair
Input unit for mismatching base pair
Search unit
407
1501
Acquisition unit for similar base sequence candidate
1401
Determination unit
1402
Apparatus for searching for base sequence
1500

Fig. 16

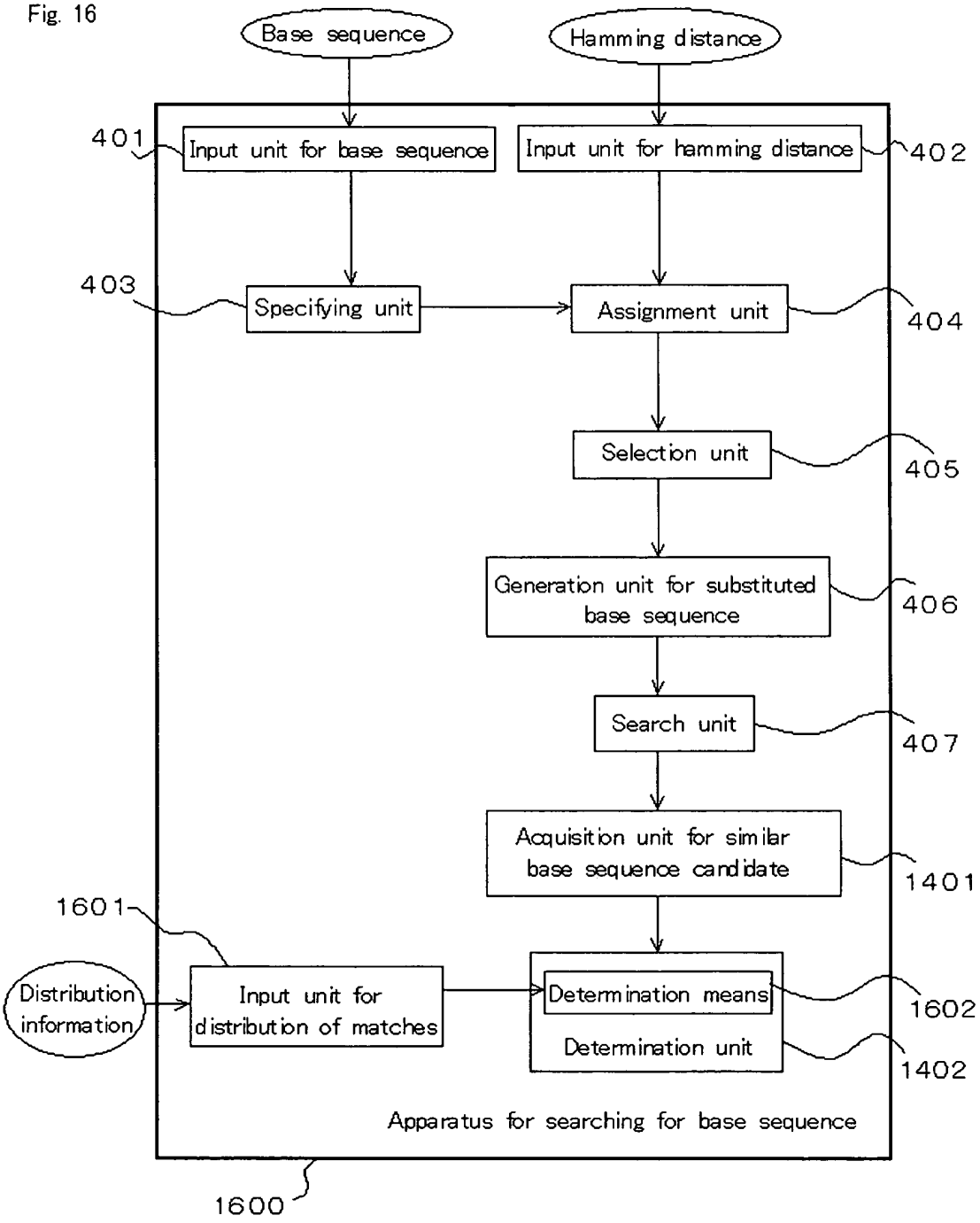

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
Search unit
407
Acquisition unit for similar base sequence candidate
1401
1601
Distribution information
Input unit for distribution of matches
Determination means
1602
Determination unit
1402
Apparatus for searching for base sequence
1600

Fig 17

```
distributeHammingDistance(int P, int H, int nSize, int* vec)
{
    int h;

    if (P==1) {
        vec[1] = h;

        /*
         * one of the assignments of the hamming distance to all partial
         * sequences is completed, so that the hamming distances
         * stored in vec are outputted
         */
        for (int i = 1; i <= nSize; i = i + 1) {
            printf("Part %d th: %d", i, vec[i]);
            /* a separator or a terminator is outputted */
            if (i != nSize) {
                /* a comma is outputted as a separator */
                printf(", ");
            }
            else {
                /* a break is outputted as a terminator */
                printf("¥n");
            }
        }

    }
    else {
        for (h = 0; h <= H; h = h + 1) {
            vec[P] = h;
            distributeHammingDistance(P - 1, H - h, nSize, vec);
        }
    }
}
```

Fig 18

```
for (I1 = 0; I1 < L; I1 = I1 + 1) {
    for (I2 = I1 + 1; I2 < L; I2 = I2 + 1) {
        foreach a1 in {A, C, G, T} {
            if (S[I1] != a1 ) {
                foreach a2 in {A, C, G, T} {
                    if (S[I2] != a2) {
                        Generate the substituted base sequence
                        by substituting I1-th base of S to a1,
                        and the I2-th base to a2
                    }
                }
            }
        }
    }
}
```

Fig 19

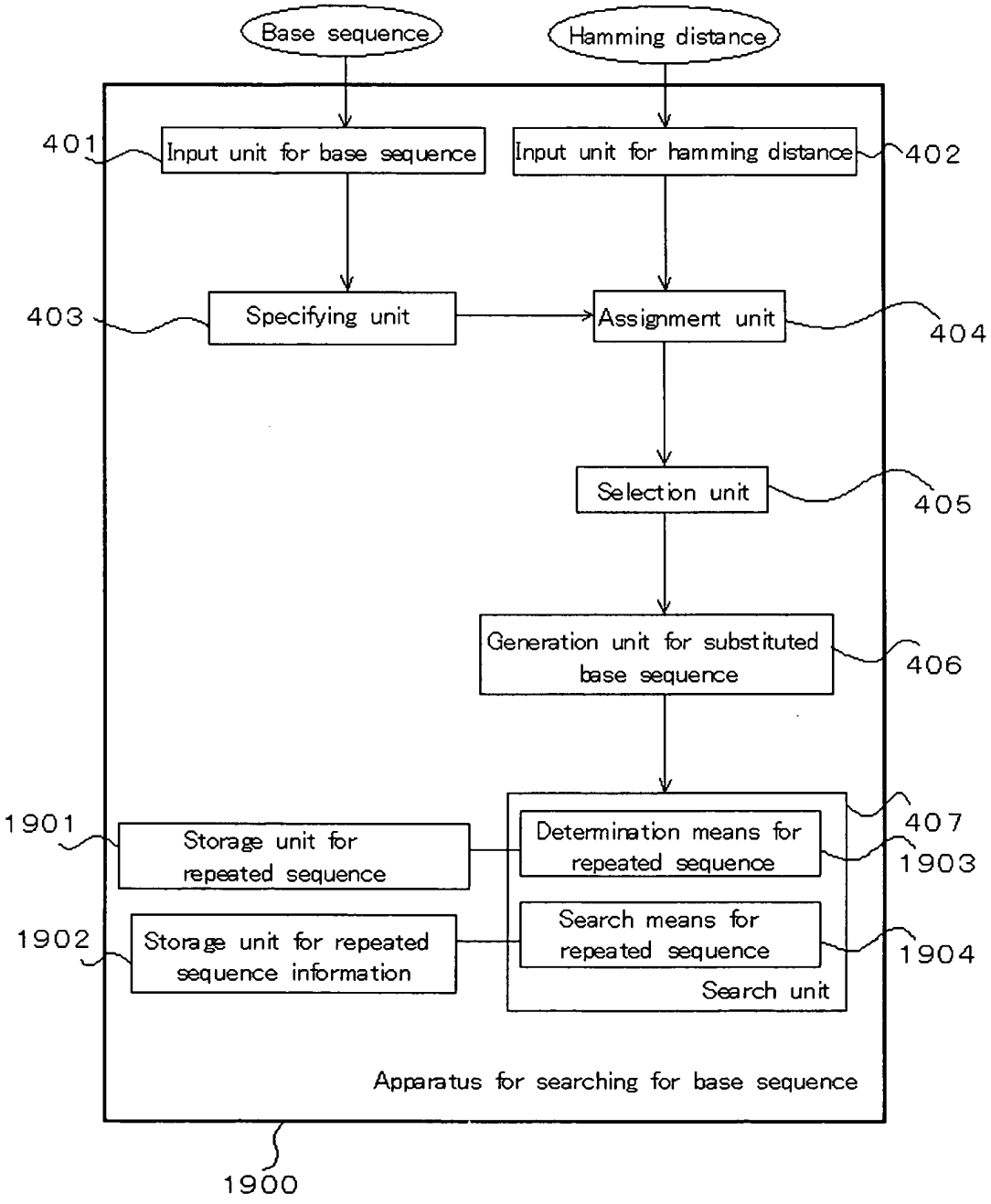

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
1901
Storage unit for repeated sequence
Determination means for repeated sequence
407
1903
1902
Storage unit for repeated sequence information
Search means for repeated sequence
1904
Search unit
Apparatus for searching for base sequence
1900

Fig. 20

Identifier of repeated sequence	Repeated sequence
1	ACGGGUC•••
2	GGGGGAAAA•••
⋮	⋮

Fig. 21

Identifier of repeated sequence	Position of appearance
1	58210
1	37703
1	27503
1	30516
⋮	⋮
2	167
2	27367
2	112109
⋮	⋮

Fig. 22

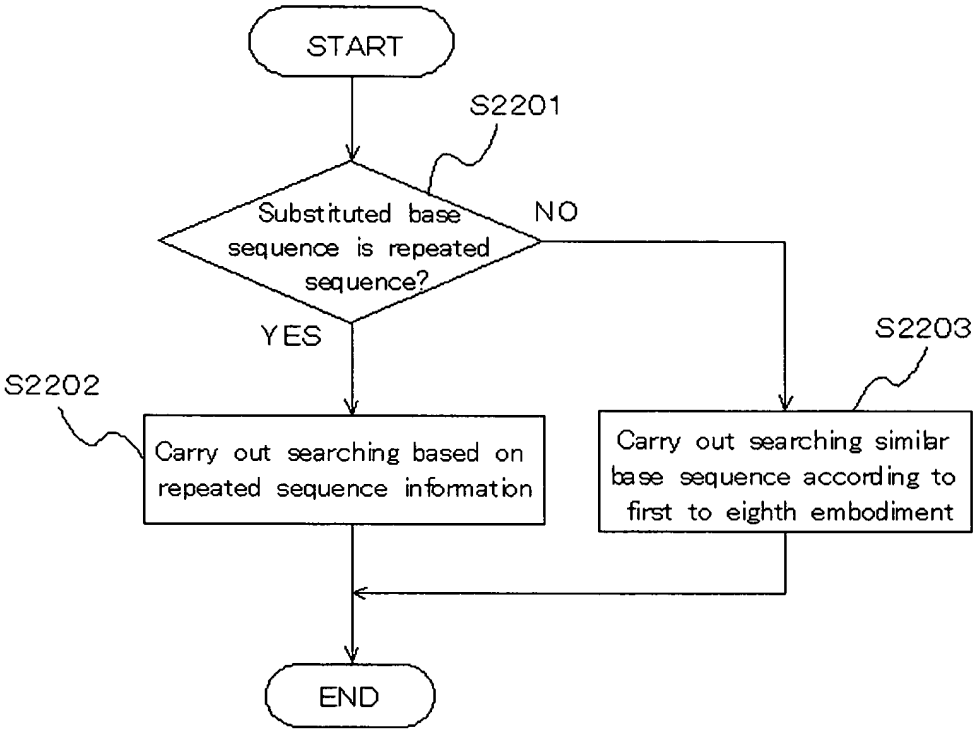

START
S2201
Substituted base sequence is repeated sequence?
NO
YES
S2202
Carry out searching based on repeated sequence information
S2203
Carry out searching similar base sequence according to first to eighth embodiment
END

Fig. 23

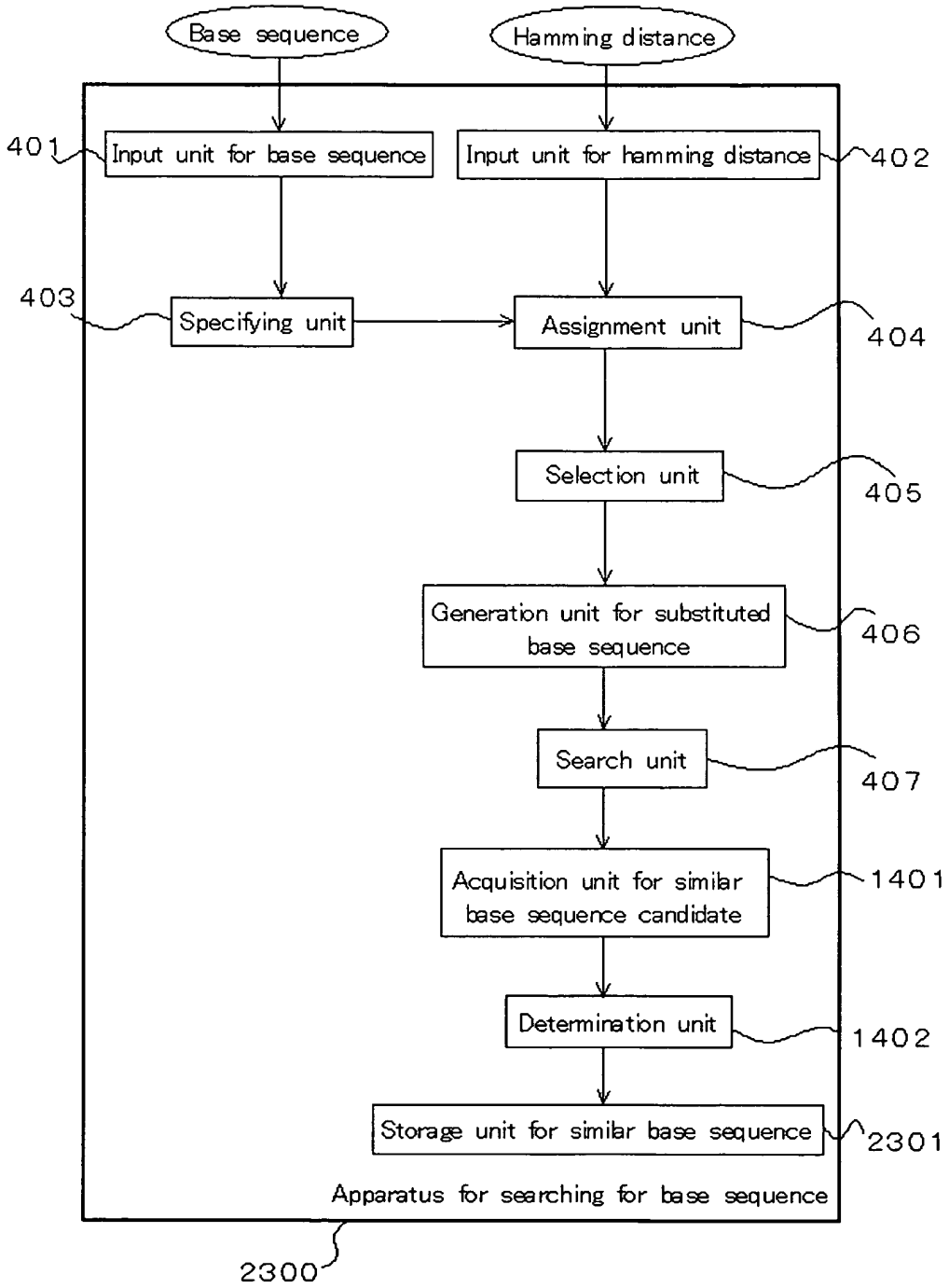

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
Search unit
407
Acquisition unit for similar base sequence candidate
1401
Determination unit
1402
Storage unit for similar base sequence
2301
Apparatus for searching for base sequence
2300

Fig. 24

Input for base sequence	Hamming distance	Similar base sequence
⋮	⋮	⋮

Fig. 25

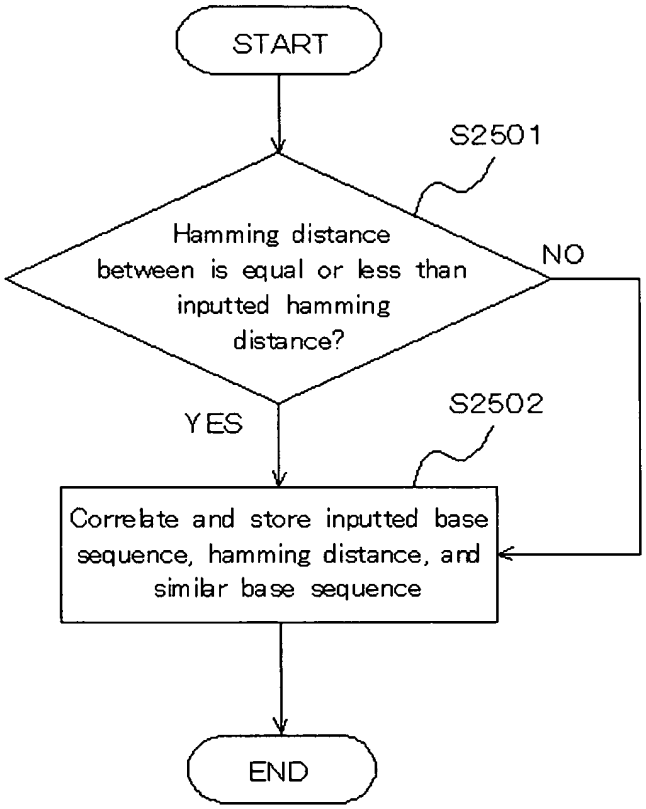

START
S2501
Hamming distance between is equal or less than inputted hamming distance?
NO
YES
S2502
Correlate and store inputted base sequence, hamming distance, and similar base sequence
END

Fig 26

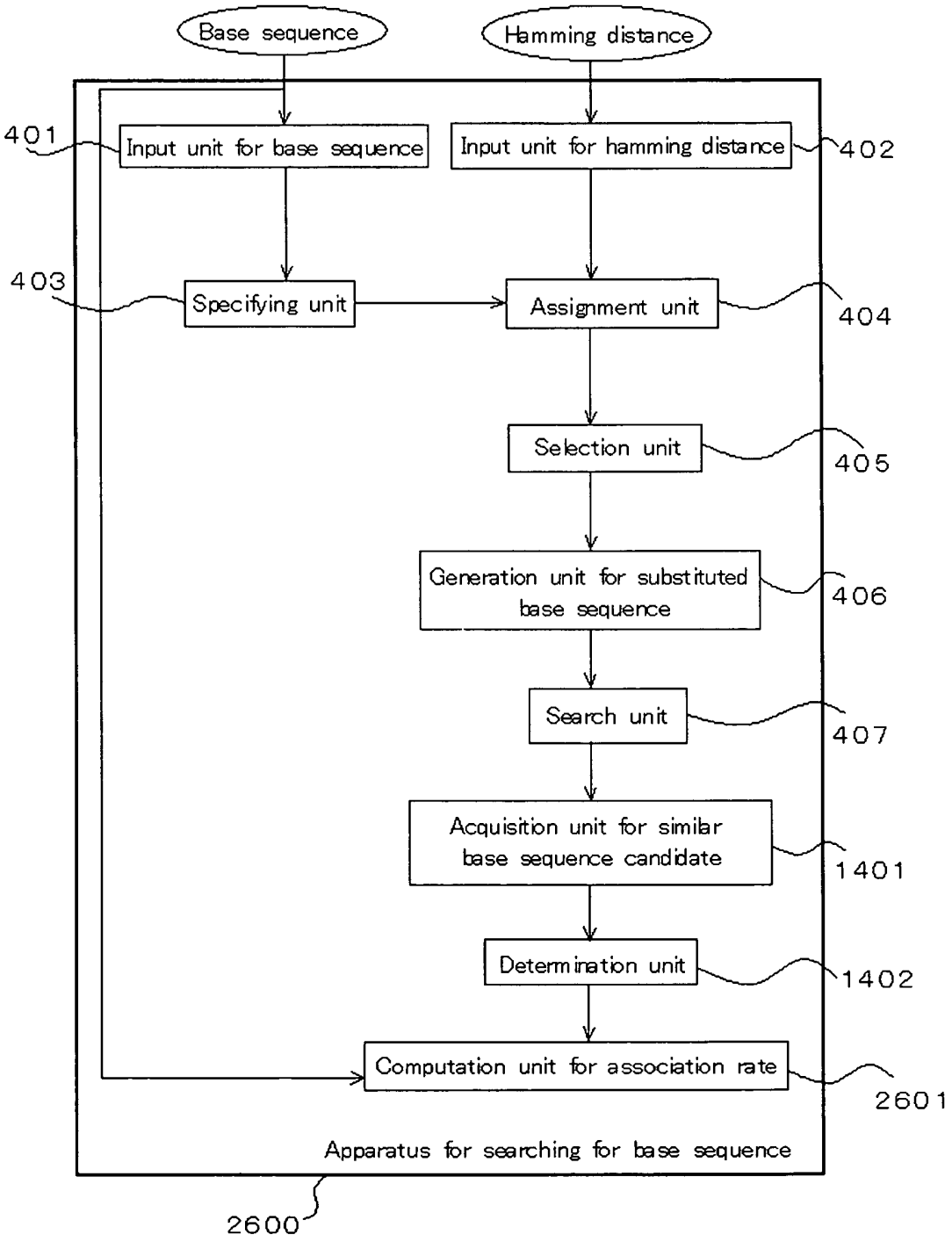

Base sequence
Hamming distance
401
Input unit for base sequence
Input unit for hamming distance
402
403
Specifying unit
Assignment unit
404
Selection unit
405
Generation unit for substituted base sequence
406
Search unit
407
Acquisition unit for similar base sequence candidate
1401
Determination unit
1402
Computation unit for association rate
2601
Apparatus for searching for base sequence
2600

Fig. 27

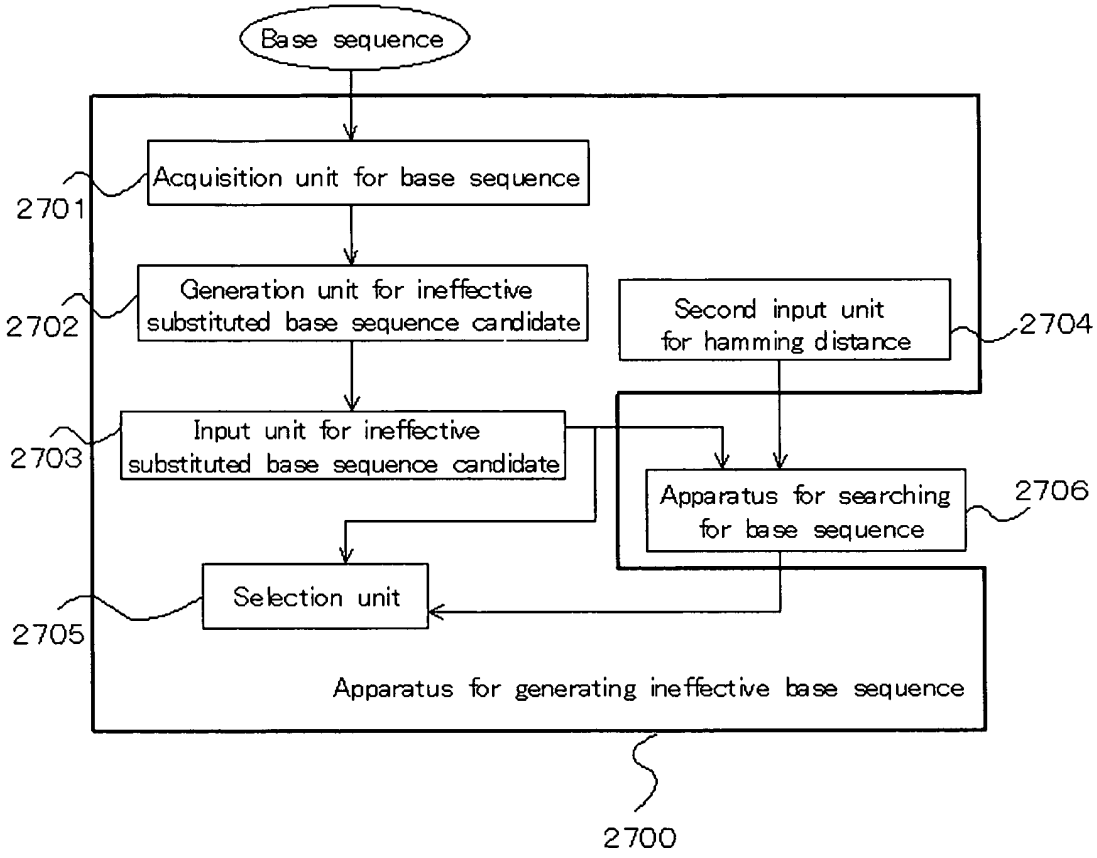

Base sequence
Acquisition unit for base sequence
2701
Generation unit for ineffective substituted base sequence candidate
2702
Second input unit for hamming distance
2704
Input unit for ineffective substituted base sequence candidate
2703
Apparatus for searching for base sequence
2706
Selection unit
2705
Apparatus for generating ineffective base sequence
2700

Fig. 28

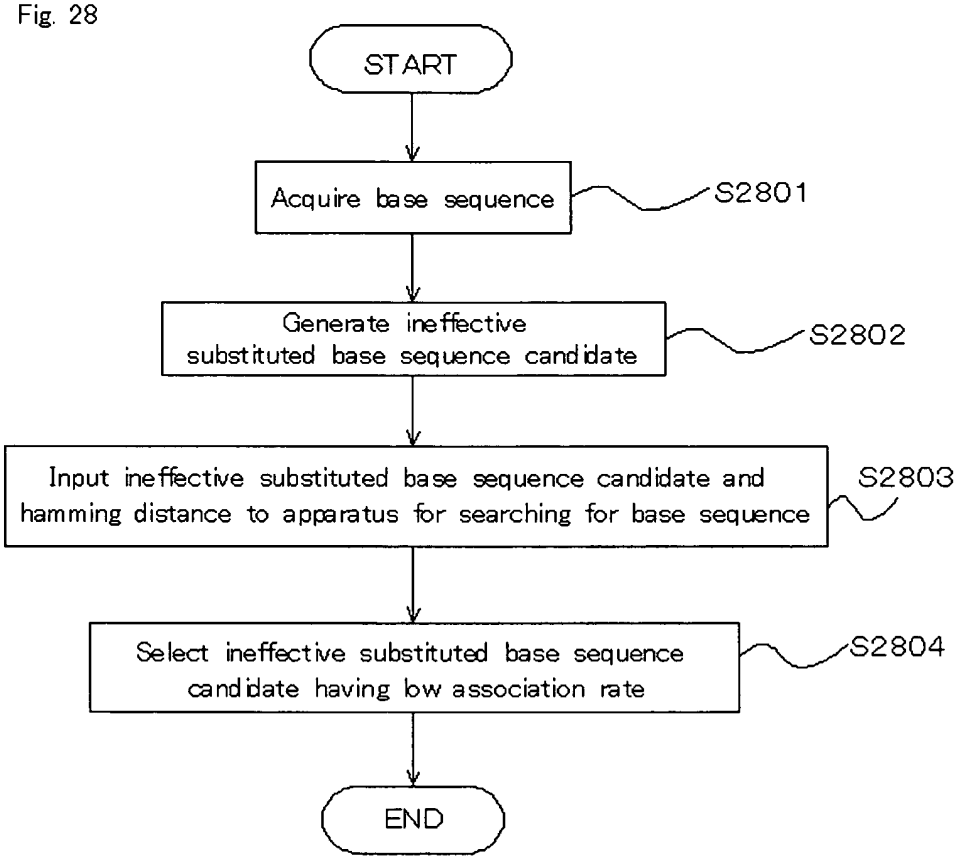

START
Acquire base sequence
S2801
Generate ineffective substituted base sequence candidate
S2802
Input ineffective substituted base sequence candidate and hamming distance to apparatus for searching for base sequence
S2803
Select ineffective substituted base sequence candidate having low association rate
S2804
END

Fig. 29

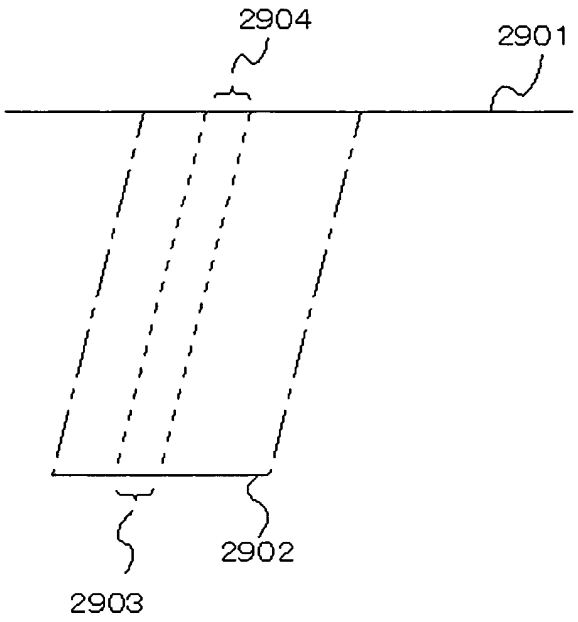


Fig. 30

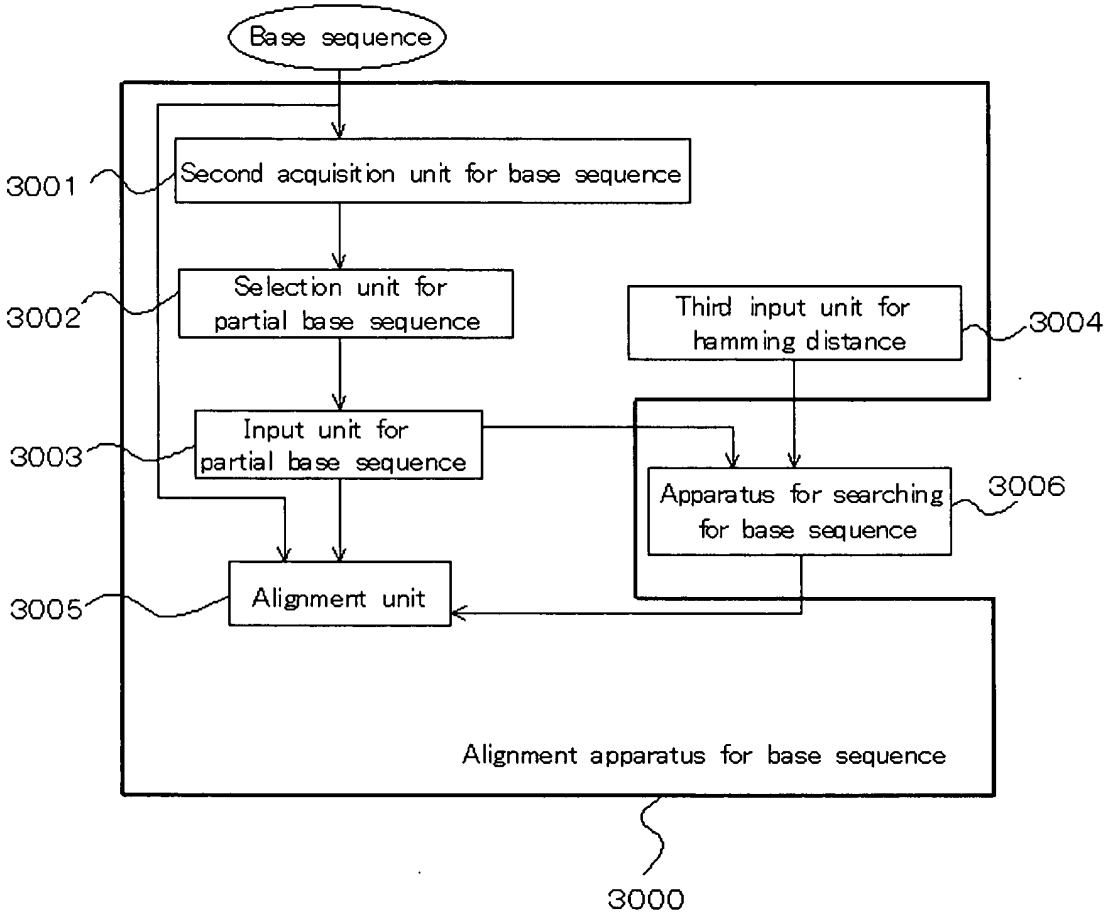
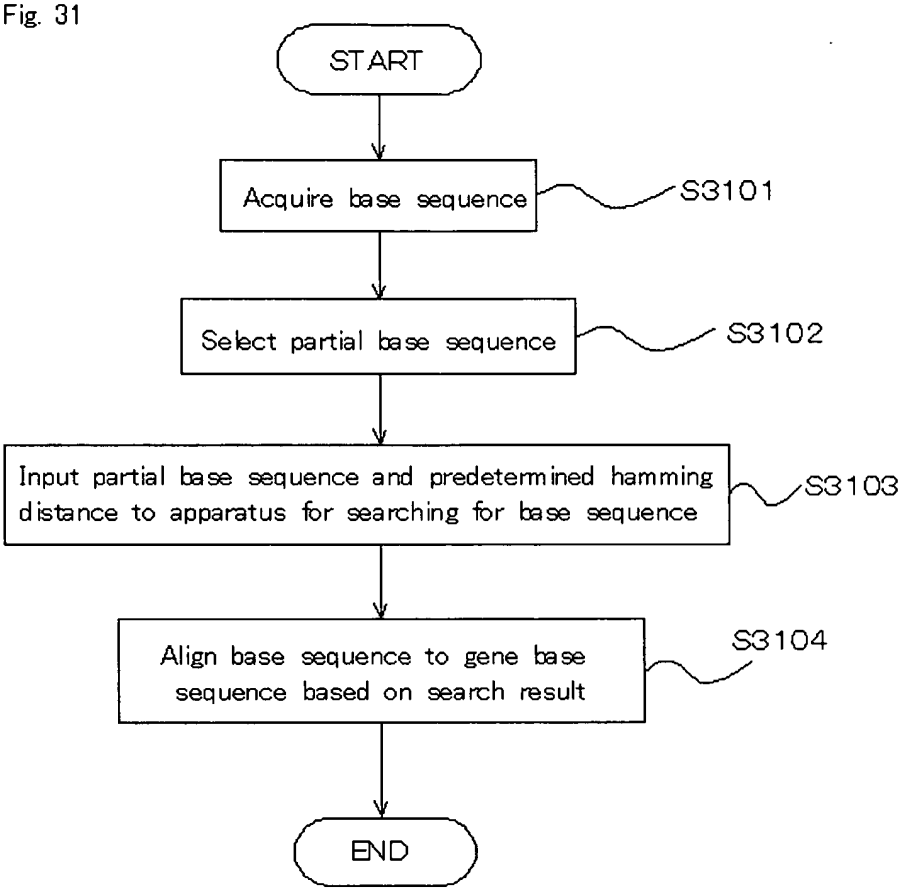

Base sequence
3001
Second acquisition unit for base sequence
3002
Selection unit for partial base sequence
Third input unit for hamming distance
3004
3003
Input unit for partial base sequence
Apparatus for searching for base sequence
3006
3005
Alignment unit
Alignment apparatus for base sequence
3000

Fig. 31


START
Acquire base sequence
S3101
Select partial base sequence
S3102
Input partial base sequence and predetermined hamming distance to apparatus for searching for base sequence
S3103
Align base sequence to gene base sequence based on search result
S3104
END